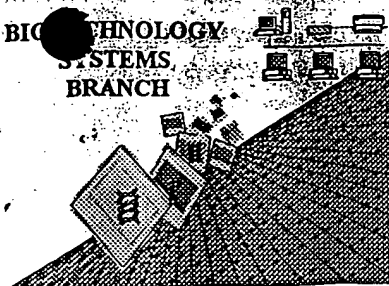


44

RAW SEQUENCE LISTING ERROR REPORT

BIO TECHNOLOGY
SYSTEMS
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/820,843

Source: OIPE

Date Processed by STIC: 4/18/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be downloaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pae/checker>

Raw Sequence Listing Error Summary

ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/820,843

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ Wrapped Nucleics The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2 ☐ Wrapped Aminos The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3 ☐ Incorrect Line Length The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4 ☐ Misaligned Amino Acid Numbering The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5 ☐ Non-ASCII This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6 ☐ Variable Length Sequence(s) ☐ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing
- 7 ☐ PatentIn ver. 2.0 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) ☐. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 8 ☐ Skipped Sequences (OLD RULES) Sequence(s) ☐ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9 ☐ Skipped Sequences (NEW RULES) Sequence(s) ☐ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 10 ☐ Use of n's or Xaa's (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11 ☐ Use of "Artificial" (NEW RULES) Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.
Valid response is Artificial Sequence.
- 12 ☐ Use of <220>Feature (NEW RULES) Sequence(s) ☐ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13 ☐ PatentIn ver. 2.0 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.

OIPE



RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/820,843

DATE: 04/18/2001
TIME: 07:35:07

Does Not Comply
Corrected Diskette Needed

Input Set : A:\ES.txt
Output Set: N:\CRF3\04182001\I820843.raw

see p. 6, too

110> APPLICANT: Council of Scientific and Industrial Research
5 <120> TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEINS
6 USEFUL AS ANTI-INFECTIVES
8 <130> FILE REFERENCE: Q63915
10 <140> CURRENT APPLICATION NUMBER: US/09/820,843
10 <141> CURRENT FILING DATE: 2001-03-30
10 <160> NUMBER OF SEQ ID NOS: 118
12 <170> SOFTWARE: PatentIn version 3.0
14 <210> SEQ ID NO: 1
15 <211> LENGTH: 51
16 <212> TYPE: PRT
17 <213> ORGANISM: C. jejuni
19 <220> FEATURE:
4--> 20 <221> NAME/KEY: misc-feature *misc-feature*
21 <223> OTHER INFORMATION: highly acidic protein
24 <220> FEATURE:
4--> 25 <221> NAME/KEY: misc-feature
26 <223> OTHER INFORMATION: gi|6967728
29 <400> SEQUENCE: 1
31 Met Ala Tyr Glu Asp Glu Glu Asp Leu Asn Tyr Asp Asp Tyr Glu Asn
32 1 5 10 15
34 Glu Asp Glu Glu Tyr Pro Gln Asn His His Lys Asn Tyr Asn Tyr Asp
35 20 25 30
37 Asp Asp Asp Tyr Glu Tyr Asp Asp Asp Asn Asn Asp Asp Asp Phe Tyr
38 35 40 45
40 Glu Met Asp
41 50
43 <210> SEQ ID NO: 2
44 <211> LENGTH: 32
45 <212> TYPE: PRT
46 <213> ORGANISM: C. jejuni
48 <220> FEATURE:
49 <221> NAME/KEY: misc_feature
50 <223> OTHER INFORMATION: small hydrophobic protein
53 <220> FEATURE:
54 <221> NAME/KEY: misc_feature
55 <223> OTHER INFORMATION: gi|6969129
58 <400> SEQUENCE: 2
60 Met Thr Met Leu Asp Ile Phe Glu Ile Ile Phe Ile Thr Thr Val Val
61 1 5 10 15
63 Ile Ile Gly Phe Gly Gly Ile Val Phe Val Val Thr Lys Glu Lys Lys
64 20 25 30
66 <210> SEQ ID NO: 3
67 <211> LENGTH: 57
68 <212> TYPE: PRT
69 <213> ORGANISM: C. jejuni
71 <220> FEATURE:

use underscore, not hyphen

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/820,843

DATE: 04/18/2001
 TIME: 07:35:07

Input Set : A:\ES.txt
 Output Set: N:\CRF3\04182001\I820843.raw

```

72 <221> NAME/KEY: misc_feature
73 <223> OTHER INFORMATION: putative coiled coil protein
76 <220> FEATURE:
77 <221> NAME/KEY: misc_feature
78 <223> OTHER INFORMATION: gi|6968493
81 <400> SEQUENCE: 3
83 Met Ser Phe Glu Glu Asn Leu Lys His Ala Asn Glu Ser Leu Glu Lys
84 1          5          10          15
86 Leu Asn Asn Gln Glu Leu Ala Leu Asp Glu Ser Val Lys Ile Tyr Lys
87          20          25          30
89 Glu Gly Leu Glu Ser Ile Lys Lys Ala Arg Leu Glu Leu Glu Lys Ala
90          35          40          45
92 Lys Leu Glu Val Glu Gln Ile Asp Glu
93          50          55
95 <210> SEQ ID NO: 4
96 <211> LENGTH: 542
97 <212> TYPE: PRT
98 <213> ORGANISM: C. jejuni
100 <220> FEATURE:
101 <221> NAME/KEY: misc_feature
102 <223> OTHER INFORMATION: highly acidic protein
105 <220> FEATURE:
106 <221> NAME/KEY: misc_feature
107 <223> OTHER INFORMATION: gi|6968611
110 <400> SEQUENCE: 4
112 Met Lys Ile Leu Leu Leu Asn Glu Asn Pro Val Val Ser Arg Leu Val
113 1          5          10          15
115 Ser Leu Ser Ala Lys Lys Met Ser Tyr Asp Phe Glu Glu Leu Asn Ala
116          20          25          30
118 Tyr Ser Glu Asn Leu Gly Asn Tyr Asp Val Ile Val Val Asp Ser Asp
119          35          40          45
121 Thr Pro Ala Pro Leu Lys Ile Leu Lys Glu Lys Cys Asp Arg Leu Ile
122          50          55          60
124 Phe Leu Ala Pro Arg Asn Gln Asn Val Glu Asp Ile Asp Ala Gln Ile
125 65          70          75          80
127 Leu Gln Lys Pro Phe Leu Pro Thr Asp Phe Leu Asn Leu Leu Asn Asn
128          85          90          95
130 Lys Asp Ala Asn Lys His Thr Ser Ile Asp Leu Pro Met Leu Ser Asn
131          100         105         110
133 Asp Glu Asn Pro Tyr Ala Asp Ile Ser Leu Asp Leu Asp Asn Leu Asn
134          115         120         125
136 Leu Asp Asp Leu Pro Asp Glu Asn Ser Leu Asp Ile Asn Ser Glu Gly
137          130         135         140
139 Met Glu Asp Leu Ser Phe Asp Asp Asp Ala Gln Asp Asp Asn Ala Asn
140 145          150         155         160
142 Lys Thr Leu Glu Thr Gln Asn Leu Glu His Glu Thr Ile Lys Glu Gln
143          165         170         175
145 Thr Gln Glu Asp Thr Gln Ile Asp Leu Asp Leu Thr Leu Glu Asp Gly
146          180         185         190

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/820,843

DATE: 04/18/2001

TIME: 07:35:07

Input Set : A:\ES.txt

Output Set: N:\CRF3\04182001\I820843.raw

```

148 Glu Ser Glu Lys Glu Asp Leu Ser Gln Glu His Thr Ala Leu Asp Thr
149      195      200      205
151 Glu Pro Ser Leu Asp Glu Leu Asp Asp Lys Asn Asp Glu Asp Leu Glu
152      210      215      220
154 Ile Lys Glu Asp Asp Lys Asn Glu Glu Ile Glu Lys Gln Glu Leu Leu
155 225      230      235      240
157 Asp Asp Ser Lys Thr Asn Thr Leu Glu Met Gln Glu Glu Leu Ser Glu
158      245      250      255
160 Ser Gln Asp Asp Asn Ser Asn Lys Thr Leu Glu Thr Gln Asn Leu Glu
161      260      265      270
163 His Asp Asn Leu Glu Gln Glu Thr Ile Lys Glu Gln Thr Gln Glu Asp
164      275      280      285
166 Thr Gln Ile Asp Leu Asp Leu Thr Leu Glu Asp Gly Glu Ser Glu Lys
167      290      295      300
169 Glu Asp Leu Ser Gln Glu His Thr Ala Leu Asp Thr Glu Pro Ser Leu
170 305      310      315      320
172 Asp Glu Leu Asp Asp Lys Asn Asp Glu Asp Leu Glu Asp Asn Lys Glu
173      325      330      335
175 Leu Gln Ala Asn Ile Ser Asp Phe Asp Asp Leu Pro Glu Val Glu Glu
176      340      345      350
178 Gln Glu Lys Glu Met Asp Phe Asp Asp Leu Pro Glu Asp Ala Glu Phe
179      355      360      365
181 Leu Gly Gln Ala Lys Tyr Asn Glu Glu Ser Glu Glu Asn Leu Glu Glu
182      370      375      380
184 Phe Ala Pro Val Val Glu Glu Asp Ile Gln Asp Glu Ile Asp Asp Phe
185 385      390      395      400
187 Ala Ser Asn Leu Ser Thr Gln Asp Gln Ile Lys Glu Glu Leu Ala Gln
188      405      410      415
190 Leu Asp Glu Leu Asp Tyr Gly Ile Asp Ser Asp Asn Ser Ser Lys Val
191      420      425      430
193 Leu Glu Asp Phe Lys Asp Glu Pro Ile Leu Asp Asp Lys Glu Leu Gly
194      435      440      445
196 Thr Asn Glu Glu Glu Val Val Val Pro Asn Leu Asn Ile Ser Asp Phe
197      450      455      460
199 Asp Thr Leu Lys Glu Ser Asp Ile Gln Glu Ala Leu Gly Glu Glu Ile
200 465      470      475      480
202 Leu Glu Lys Asn Glu Glu Pro Ile Val Ser Asp Val Thr Lys Asp Asp
203      485      490      495
205 Asn Ser Glu Glu Ile Val Asn Glu Leu Ser Gln Ser Ile Ala Gly Ala
206      500      505      510
208 Ile Thr Ser Ser Ile Lys Asp Asp Thr Leu Lys Ala Ala Leu Lys Gly
209      515      520      525
211 Met Asn Met Asn Ile Asn Ile Asn Ile Ser Phe Lys Glu Asp
212      530      535      540
214 <210> SEQ ID NO: 5
215 <211> LENGTH: 172
216 <212> TYPE: PRT
217 <213> ORGANISM: C. pneumoniaeCWL029
219 <220> FEATURE:

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/820,843
 DATE: 04/18/2001
 TIME: 07:35:07

Input Set : A:\ES.txt
 Output Set: N:\CRF3\04182001\I820843.raw

220 <221> NAME/KEY: misc_feature
 221 <223> OTHER INFORMATION: histone like protein 2
 224 <220> FEATURE:
 225 <221> NAME/KEY: misc_feature
 226 <223> OTHER INFORMATION: gi|4376663
 229 <400> SEQUENCE: 5
 231 Met Ile Gly Ala Gln Lys Lys Gln Ser Gly Lys Lys Thr Ala Ser Arg
 232 1 5 10 15
 234 Ala Val Arg Lys Pro Ala Lys Lys Val Ala Ala Lys Arg Thr Val Lys
 235 20 25 30
 237 Lys Ala Thr Val Arg Lys Thr Ala Val Lys Lys Pro Ala Val Arg Lys
 238 35 40 45
 240 Thr Ala Ala Lys Lys Thr Val Ala Lys Lys Thr Thr Ala Lys Arg Thr
 241 50 55 60
 243 Val Arg Lys Thr Val Ala Lys Lys Pro Ala Val Lys Lys Val Ala Ala
 244 65 70 75 80
 246 Lys Arg Val Val Lys Lys Thr Val Ala Lys Lys Thr Thr Ala Lys Arg
 247 85 90 95
 249 Ala Val Arg Lys Thr Val Ala Lys Lys Pro Val Ala Arg Lys Thr Thr
 250 100 105 110
 252 Val Ala Lys Gly Ser Pro Lys Lys Ala Ala Ala Cys Ala Leu Ala Cys
 253 115 120 125
 255 His Lys Asn His Lys His Thr Ser Ser Cys Lys Arg Val Cys Ser Ser
 256 130 135 140
 258 Thr Ala Thr Arg Lys His Gly Ser Lys Ser Arg Val Arg Thr Ala His
 259 145 150 155 160
 261 Gly Trp Arg His Gln Leu Ile Lys Met Met Ser Arg
 262 165 170
 264 <210> SEQ ID NO: 6
 265 <211> LENGTH: 63
 266 <212> TYPE: PRT
 267 <213> ORGANISM: C. trachomatis
 269 <220> FEATURE:
 270 <221> NAME/KEY: misc_feature
 271 <223> OTHER INFORMATION: hypothetical protein-possible frameshift with CT593
 274 <220> FEATURE:
 275 <221> NAME/KEY: misc_feature
 276 <223> OTHER INFORMATION: gi|3522902
 279 <400> SEQUENCE: 6
 281 Met Phe Thr Leu Phe Leu Cys Glu His Leu Leu Thr Asn Ile Leu Ala
 282 1 5 10 15
 284 Ser Ser Phe Leu Ala Lys Ser Gln Gly Phe Ile Thr Leu Val Asn Leu
 285 20 25 30
 287 Phe His Lys Ile Pro Gly Leu Lys Val Ile Glu Ile Thr Cys Leu Ala
 288 35 40 45
 290 Leu Pro Leu Gly Ile His Ser Ile Ile Gly Phe Ser Tyr Leu Leu
 291 50 55 60
 293 <210> SEQ ID NO: 7
 294 <211> LENGTH: 203

<210> 32
 <211> 497
 <212> PRT
 <213> B. burgdorferi

<220>
 <221> misc_feature
 <223> predicted coding region BB0553

<220>
 <221> misc_feature
 <223> gi|2688482

<400> 32

Met Asn Lys Thr Lys Asn Arg Ser Leu Thr Tyr Phe Ile Ile Leu Ser
 1 5 10 15

Cys Ile Ser Leu Phe Gly Ala Asn Asn Asn Thr Ile Ser Tyr Ser Ser
 20 25 30

Ile Glu Ile Pro Leu Glu Asp Leu Ser Glu Glu Phe Lys Ser Ser Gly
 35 40 45

Asn Lys Ser Asp Gln Ile Asn Thr Ser Lys His Leu Asn Lys Asn Ile
 50 55 60

Val Ser Tyr Glu Asp Pro Lys Lys Gly Lys Asp Leu Lys Leu Pro Glu
 65 70 75 80

Asn Ile Arg Asp Lys Lys Leu Pro Gln Lys Arg Met Asp Glu Asn Asp
 85 90 95

Leu Lys Ser Val Ile Glu Asn Tyr Glu Asn Lys Ile Lys Asn Ile Glu
 100 105 110

Lys Leu Leu Lys Thr Lys Asn Gln Lys Thr Ser Glu Asn Glu Asn Lys
 115 120 125

Lys Ile Glu Ser Ile Glu Lys Lys Ala Lys Lys Tyr Glu Ile Leu Thr
 130 135 140

Asn Lys Leu Lys Asn Glu Ile Val Glu Ile Lys Lys Leu Leu Asn Lys
 145 150 155 160

Lys Ile Lys Pro Lys Glu Asp Glu Asn Tyr Glu Lys Ile Asn Ile Glu
 165 170 175

Asn Ile Glu Glu Glu Thr Asp Asp Asp Phe Glu Asp Asn Tyr Glu Tyr
 180 185 190

Asn Asp Glu Ile Glu Xaa Thr Asn Glu Asp Asn Tyr Pro Ser Asn Glu
 195 200 205

Gly Ile Ile Asn Asn Leu Lys Glu Asn Leu Asn Glu Asn Glu Lys Tyr

*only the
 error
 portion of
 sequence 32
 shown*

*see item 10
 on Error
 Summary
 sheet*

Please Note:

Use f n and/ r Xaa have been detected in the Sequenc Listing. Please review the Sequenc Listing to nsure that a corresponding explanation is presented in th <220> to <223> fields f each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/820,843

DATE: 04/18/2001

TIME: 07:35:08

Input Set : A:\ES.txt

Output Set: N:\CRF3\04182001\I820843.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:20 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:25 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:2886 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:32
L:2886 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:4688 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:66
L:4688 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66



#4 Seebon

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Samir Kumar BRAHMACHARI et al.

Appln. No.: 09/820,843

Group Art Unit: Not Yet Assigned

Filed: March 30, 2001

Examiner: Not Yet Assigned

Confirmation number: 7045

For: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
PROTEINS USEFUL AS ANTI-INFECTIVES

**STATEMENT TO SUPPORT FILING AND SUBMISSION IN
ACCORDANCE WITH 37 C.F.R. §§ 1.821-1.825**

Assistant Commissioner for Patents
Washington, D.C. 20231
Box SEQUENCE

Sir:

In connection with a Sequence Listing submitted concurrently herewith, the undersigned
hereby states that:

1. the submission, filed herewith in accordance with 37 C.F.R. § 1.821(g), does not
include any new matter;
2. the content of the attached paper copy and the attached computer readable copy of
the Sequence Listing, submitted in accordance with 37 C.F.R. § 1.821(c) and (e), respectively,
are the same; and
3. all statements made herein of my own knowledge are true and that all statements
made on information and belief are believed to be true, and further, that these statements were
made with the knowledge that willful false statements and the like so made are punishable by

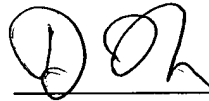
09820843-051801

**STATEMENT TO SUPPORT FILING AND SUBMISSION
IN ACCORDANCE WITH 37 C.F.R. §§ 1.821-1.825**

Q63915

fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent resulting therefrom.

Respectfully submitted,



Drew Hissong
Registration No. 44,765

SUGHRUE, MION, ZINN,
MACPEAK & SEAS, PLLC
2100 Pennsylvania Avenue, N.W.
Washington, D.C. 20037-3213
Telephone: (202) 293-7060
Facsimile: (202) 293-7860

Date: May 18, 2001

09220243-051204
T02150-ET022860